

The Digital Humanities cycle: hermeneutics, heuristics, and source criticism in a digital age

Jesper Verhoef
Utrecht University
j.verhoef@uu.nl

Melvin Wevers
Utrecht University
m.j.h.f.wevers@uu.nl

This paper demonstrates how we have applied a question-driven approach to digital history using computational methods. We use the analogy of the Digital Humanities Cycle—an updated version of the empirical cycle—to explain the iterative process of heuristics, hermeneutics, tool criticism, corpus faceting, and source criticism. The Digital Humanities Cycle entails that, while doing historical research, the strength of numerous digital tools and archives should constantly be combined and alternated with ‘classical’ historical hermeneutics, i.e. source criticism and interpretation. Figure 1 shows a visualization of the way in which the Cycle functions.

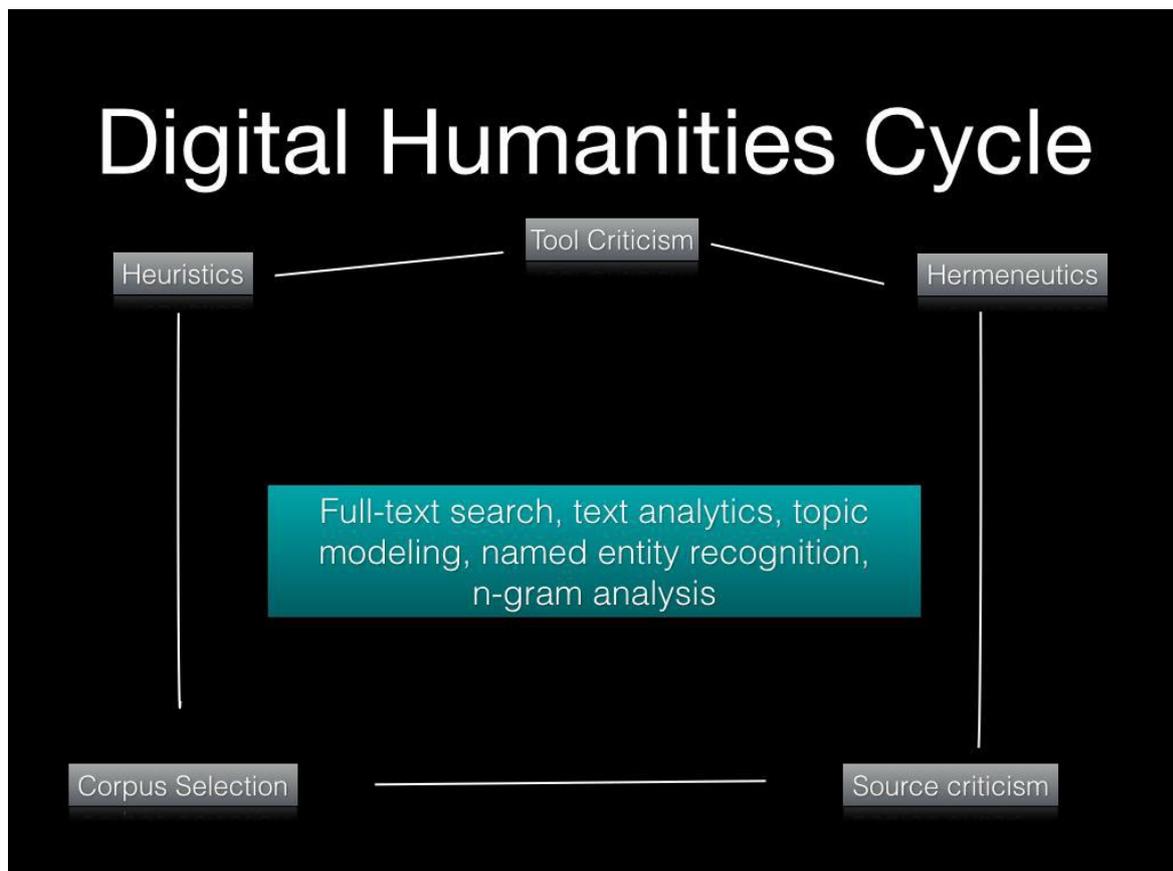


Figure 1: The Digital Humanities Cycle

We will briefly touch upon each of these aspects through examples from our own research that deal with the representation of America in Dutch digitized newspapers. These examples include debates on transistor radios and the construction of geographic connotations of consumer goods.

In our research, we have applied the Cycle to the digitized repository of the National Library of the Netherlands, which contains over 600.000 digitized newspapers between 1890 and 1990. The paper describes how full-text search, N-gram analysis, Topic Modeling, and Named Entity Recognition help to select (sub-)corpora of sources, and generate themes and topics of interest to the study. More importantly, we will show that the stages in the Cycle – searching the archive (heuristics), understanding whether and how a tool works, confining a corpus of articles and applying corpus linguistics and source criticism, and interpreting the results (hermeneutics) – are all interrelated. For instance, when applying topic modeling on all the newspapers articles that contain the word ‘portable radio’ – a result of full text search in the first place –, the topics which are produced demand interpretation. Moreover, they render new (full text) search words, and at the same time – when topics clearly demarcate a single theme – may lead a researcher to apply other digital tools to this very dataset. This could include corpus linguistic tools to clean up the corpus or extract other meaningful linguistic information. Put differently, the output of one single tool is never the end-point of research, but one of the many step needed to weave a historical narrative.

The Digital Humanities Cycle is a response to the plea made by numerous scholars that, after the Digital or Computation Turn (Nicholson 2013; Berry 2011), one should obtain ‘hybridity’, that is combine classic and digital historical research (Zaagsma 2013). When applied in combination, these techniques might yield patterns in the dataset, without missing “the power of the particular” (Hitchcock 2014). We will show that the Digital Humanities Cycle leads to a fruitful cooperation between “counting and understanding” (Rieder & Röhle 2012) – one of the prominent challenges when using digital tools.

References

- Berry D.M. (2014), ‘The computational turn: thinking about the digital humanities’, *Culture Machine*, vol. 12, pp. 1–22.
- Hitchcock T. (2014), ‘Big data, small data and meaning’ on *Historyonics* (Blog). Available from: <http://historyonics.blogspot.co.uk/2014/11/big-data-small-data-and-meaning_9.html> [13 November 2014].
- Nicholson B. (2013), ‘The digital turn’, *Media History*, vol. 19, no. 1, pp. 59–73.
- Rieder, B. and Röhle T. (2012), ‘Digital methods: five challenges’ in *Understanding Digital Humanities* ed. D.M. Berry, Houndmills: Palgrave Macmillan, pp. 67-84.
- Zaagsma Z. (2013), ‘On digital history’, *BMGN - Low Countries Historical Review* vol. 128, no. 4. pp.3-29.